

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: NI, JIAN
ROSEN, CRAIG A.
PAN, JAMES G.
GENTZ, REINER L.
DIXIT, VISHVA M.
- (ii) TITLE OF INVENTION: Death Domain Containing Receptor-4
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
(B) STREET: 9410 KEY WEST AVENUE
(C) CITY: ROCKVILLE
(D) STATE: MD
(E) COUNTRY: US
(F) ZIP: 20850
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE: 28-JAN-1997
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: BROOKES, ANDERS A
(B) REGISTRATION NUMBER: 36,373
(C) REFERENCE/DOCKET NUMBER: PF355
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (301) 309-8504
(B) TELEFAX: (301) 309-8512

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2152 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 19..1422

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTGGGGCACG AGGGCAGG ATG GCG CCA CCA CCA GCT AGA GTA CAT CTA GGT	51
Met Ala Pro Pro Pro Ala Arg Val His Leu Gly	
1 5 10	
GCG TTC CTG GCA GTG ACT CCG AAT CCC GGG AGC GCA / GCG AGT GGG ACA	99
Ala Phe Leu Ala Val Thr Pro Asn Pro Gly Ser Ala Ala Ser Gly Thr	
15 20 25	
GAG GCA GCC GCG GCC ACA CCC AGC AAA GTG TGG GGC' TCT TCC GCG GGG	147
Glu Ala Ala Ala Ala Thr Pro Ser Lys Val Trp Gly Ser Ser Ala Gly	
30 35 40	
AGG ATT GAA CCA CGA GGC GGG GGC CGA GGA GCG CTC CCT ACC TCC ATG	195
Arg Ile Glu Pro Arg Gly Gly Arg Gly Ala Leu Pro Thr Ser Met	
45 50 55	
GGA CAG CAC GGA CCC AGT GCC CGG GCC CGG GCA GGG CGC GCC CCA GGA	243
Gly Gln His Gly Pro Ser Ala Arg Ala Arg Ala Gly Arg Ala Pro Gly	
60 65 70 75	
CCC AGG CCG GCG CGG GAA GCC AGC CCT CGG CTC CGG GTC CAC AAG ACC	291
Pro Arg Pro Ala Arg Glu Ala Ser Pro Arg Leu Arg Val His Lys Thr	
80 85 90	
TTC AAG TTT GTC GTC GGG GTC CTG CTG CAG GTC GTA CCT AGC TCA	339
Phe Lys Phe Val Val Val Gly Val Leu Leu Gln Val Val Pro Ser Ser	
95 100 105	
GCT GCA ACC ATC AAA CTT CAT GAT CAA TCA ATT GGC ACA CAG CAA TGG	387
Ala Ala Thr Ile Lys Leu His Asp Gln Ser Ile Gly Thr Gln Gln Trp	
110 115 120	
GAA CAT AGC CCT TTG GGA GAG TTG TGT CCA CCA GGA TCT CAT AGA TCA	435
Glu His Ser Pro Leu Gly Glu Leu Cys Pro Pro Gly Ser His Arg Ser	
125 130 135	
GAA CGT CCT GGA GCC TGT AAC CGG TGC ACA GAG GGT GTG GGT TAC ACC	483
Glu Arg Pro Gly Ala Cys Asn Arg Cys Thr Glu Gly Val Gly Tyr Thr	
140 145 150 155	
AAT GCT TCC AAC AAT TTG TTT GCT TGC CTC CCA TGT ACA GCT TGT AAA	531
Asn Ala Ser Asn Asn Leu Phe Ala Cys Leu Pro Cys Thr Ala Cys Lys	
160 165 170	
TCA GAT GAA GAA GAG AGA AGT CCC TGC ACC ACG ACC AGG AAC ACA GCA	579
Ser Asp Glu Glu Glu Arg Ser Pro Cys Thr Thr Arg Asn Thr Ala	
175 180 185	

TGT CAG TGC AAA CCA GGA ACT TTC CGG AAT GAC AAT TCT GCT GAG ATG Cys Gln Cys Lys Pro Gly Thr Phe Arg Asn Asn Ser Ala Glu Met 190 195 200	627
TGC CGG AAG TGC AGC ACA GGG TGC CCC AGA GGG ATG GTC AAG GTC AAG Cys Arg Lys Cys Ser Thr Gly Cys Pro Arg Gly Met Val Lys Val Lys 205 210 215	675
GAT TGT ACG CCC TGG AGT GAC ATC GAG TGT GTC CAC AAA GAA TCA GGC Asp Cys Thr Pro Trp Ser Asp Ile Glu Cys Val His Lys Glu Ser Gly 220 225 230 235	723
AAT GGA CAT AAT ATA TGG GTG ATT TTG GTT GTG ACT TTG GTT GTT CCG Asn Gly His Asn Ile Trp Val Ile Leu Val Val Thr Leu Val Val Pro 240 245 250	771
TTG CTG TTG GTG GCT GTG CTG ATT GTC TGT TGC ATC GGC TCA GGT Leu Leu Leu Val Ala Val Leu Ile Val Cys Cys Cys Ile Gly Ser Gly 255 260 265	819
TGT GGA GGG GAC CCC AAG TGC ATG GAC AGG GTG TGT TTC TGG CGC TTG Cys Gly Gly Asp Pro Lys Cys Met Asp Arg Val Cys Phe Trp Arg Leu 270 275 280	867
GGT CTC CTA CGA GGG CCT GGG GCT GAG GAC AAT GCT CAC AAC GAG ATT Gly Leu Leu Arg Gly Pro Gly Ala Glu Asp Asn Ala His Asn Glu Ile 285 290 295	915
CTG AGC AAC GCA GAC TCG CTG TCC ACT TTC GTC TCT GAG CAG CAA ATG Leu Ser Asn Ala Asp Ser Leu Ser Thr Phe Val Ser Glu Gln Gln Met 300 305 310 315	963
GAA AGC CAG GAG CCG GCA GAT TTG ACA GGT GTC ACT GTA CAG TCC CCA Glu Ser Gln Glu Pro Ala Asp Leu Thr Gly Val Thr Val Gln Ser Pro 320 325 330	1011
GGG GAG GCA CAG TGT CTG GGA CCG GCA GAA GCT GAA GGG TCT CAG Gly Glu Ala Gln Cys Leu Leu Gly Pro Ala Glu Ala Glu Gly Ser Gln 335 340 345	1059
AGG AGG AGG CTG CTG GTT CCA GCA AAT GGT GCT GAC CCC ACT GAG ACT Arg Arg Arg Leu Leu Val Pro Ala Asn Gly Ala Asp Pro Thr Glu Thr 350 355 360	1107
CTG ATG CTG TTC TTT GAC AAG TTT GCA AAC ATC GTG CCC TTT GAC TCC Leu Met Leu Phe Phe Asp Lys Phe Ala Asn Ile Val Pro Phe Asp Ser 365 370 375	1155
TGG GAC CAG CTC ATG AGG CAG CTG GAC CTC ACG AAA AAT GAG ATC GAT Trp Asp Gln Leu Met Arg Gln Leu Asp Leu Thr Lys Asn Glu Ile Asp 380 385 390 395	1203
GTG GTC AGA GCT GGT ACA GCA GGC CCA GGG GAT GCC TTG TAT GCA ATG Val Val Arg Ala Gly Thr Ala Gly Pro Gly Asp Ala Leu Tyr Ala Met	1251

400

405

410

CTG ATG AAA TGG GTC AAC AAA ACT GGA CGG AAC GCC TCG ATC CAC ACC Leu Met Lys Trp Val Asn Lys Thr Gly Arg Asn Ala Ser Ile His Thr 415 420 425	1299
CTG CTG GAT GCC TTG GAG AGG ATG GAA GAG AGA CAT GCA AAA GAG AAG Leu Leu Asp Ala Leu Glu Arg Met Glu Glu Arg His Ala Lys Glu Lys 430 435 440	1347
ATT CAG GAC CTC TTG GTG GAC TCT GGA AAG TTC ATC TAC TTA GAA GAT Ile Gln Asp Leu Leu Val Asp Ser Gly Lys Phe Ile Tyr Leu Glu Asp 445 450 455	1395
GGC ACA GGC TCT GCC GTG TCC TTG GAG TGAAAGACTC TTTTTACCA Gly Thr Gly Ser Ala Val Ser Leu Glu 460 465	1442
AGGTTTCCTC TTAGGTGTTA GGAGTTAATA CATATTAGGT TTTTTTTTTT TTTAACATGT	1502
ATACAAAGTA AATTCTTAGC CACGTGTATT GGCTCCTGCC TGTAATCCCA TCACTTGGG	1562
AGGCTGACGC CGGTGGATCC ACTTGAGGTC CGAAGTTCCA AGACCAGCCC TGAACCAACA	1622
TCGTGGAAAT GCCCCGTCTTT TACAAAAAAA TACAAAAAAAT TCAACTGGAA TGTGCATGGT	1682
GTGTGCCATC ATTTCTCGG CTAACTACGG GAGGTCTGAG GCCAGGAGAA TCCACTTGAA	1742
CCCCACGAAG GACAGTGTAG ACTGCAGATT GCACCACTGC ACTCCCAGCC TGGGAACACA	1802
GAGCAAGACT CTGTCTCAAG ATAAAATAAA ATAAAATTGA AAGAATTATT GCCCGACTGA	1862
GGCTCACATG CCAAAGGAAA ATCTGGTTCT CCCCTGAGCT GGCTCCGTG TGTTCCTTA	1922
TCATGGTGGT CAATTGGAGG TGTTAATTG AATGGATTAA GGAACACCTA GAACACTGGT	1982
AAGGCATTAT TTCTGGGACA TTATTTCTGG GCATGTCTTC GAGGGTGTAA CCAGAGGGGA	2042
TTGGCATGCG ATCGGGTGGAA CTGAGTGGAA AAGACCTACC CTTAATTGG GGGGGCACCG	2102
TCCGACAGAC TGGGGAGCAA GATAGAAGAA AACAAAAAAA AAAAAAAA	2152

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 468 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Pro Pro Pro Ala Arg Val His Leu Gly Ala Phe Leu Ala Val

1	5	10	15
Thr Pro Asn Pro Gly Ser Ala Ala Ser Gly Thr Glu Ala Ala Ala Ala			
20	25	30	
Thr Pro Ser Lys Val Trp Gly Ser Ser Ala Gly Arg Ile Glu Pro Arg			
35	40	45	
Gly Gly Gly Arg Gly Ala Leu Pro Thr Ser Met Gly Gln His Gly Pro			
50	55	60	
Ser Ala Arg Ala Arg Ala Gly Arg Ala Pro Gly Pro Arg Pro Ala Arg			
65	70	75	80
Glu Ala Ser Pro Arg Leu Arg Val His Lys Thr Phe Lys Phe Val Val			
85	90	95	
Val Gly Val Leu Leu Gln Val Val Pro Ser Ser Ala Ala Thr Ile Lys			
100	105	110	
Leu His Asp Gln Ser Ile Gly Thr Gln Gln Trp Glu His Ser Pro Leu			
115	120	125	
Gly Glu Leu Cys Pro Pro Gly Ser His Arg Ser Glu Arg Pro Gly Ala			
130	135	140	
Cys Asn Arg Cys Thr Glu Gly Val Gly Tyr Thr Asn Ala Ser Asn Asn			
145	150	155	160
Leu Phe Ala Cys Leu Pro Cys Thr Ala Cys Lys Ser Asp Glu Glu Glu			
165	170	175	
Arg Ser Pro Cys Thr Thr Arg Asn Thr Ala Cys Gln Cys Lys Pro			
180	185	190	
Gly Thr Phe Arg Asn Asp Asn Ser Ala Glu Met Cys Arg Lys Cys Ser			
195	200	205	
Thr Gly Cys Pro Arg Gly Met Val Lys Val Lys Asp Cys Thr Pro Trp			
210	215	220	
Ser Asp Ile Glu Cys Val His Lys Glu Ser Gly Asn Gly His Asn Ile			
225	230	235	240
Trp Val Ile Leu Val Val Thr Leu Val Val Pro Leu Leu Leu Val Ala			
245	250	255	
Val Leu Ile Val Cys Cys Cys Ile Gly Ser Gly Cys Gly Gly Asp Pro			
260	265	270	
Lys Cys Met Asp Arg Val Cys Phe Trp Arg Leu Gly Leu Leu Arg Gly			
275	280	285	
Pro Gly Ala Glu Asp Asn Ala His Asn Glu Ile Leu Ser Asn Ala Asp			
290	295	300	

Ser Leu Ser Thr Phe Val Ser Glu Gln Gln Met Glu Ser Gln Glu Pro
 305 310 315 320
 Ala Asp Leu Thr Gly Val Thr Val Gln Ser Pro Gly Glu Ala Gln Cys
 325 330 335
 Leu Leu Gly Pro Ala Glu Ala Glu Gly Ser Gln Arg Arg Arg Leu Leu
 340 345 350
 Val Pro Ala Asn Gly Ala Asp Pro Thr Glu Thr Leu Met Leu Phe Phe
 355 360 365
 Asp Lys Phe Ala Asn Ile Val Pro Phe Asp Ser Trp Asp Gln Leu Met
 370 375 380
 Arg Gln Leu Asp Leu Thr Lys Asn Glu Ile Asp Val Val Arg Ala Gly
 385 390 395 400
 Thr Ala Gly Pro Gly Asp Ala Leu Tyr Ala Met Leu Met Lys Trp Val
 405 410 415
 Asn Lys Thr Gly Arg Asn Ala Ser Ile His Thr Leu Leu Asp Ala Leu
 420 425 430
 Glu Arg Met Glu Glu Arg His Ala Lys Glu Lys Ile Gln Asp Leu Leu
 435 440 445
 Val Asp Ser Gly Lys Phe Ile Tyr Leu Glu Asp Gly Thr Gly Ser Ala
 450 455 460
 Val Ser Leu Glu
 465

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 669 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Leu Gly Ile Trp Thr Leu Leu Pro Leu Val Leu Thr Ser Val Ala
 1 5 10 15
 Arg Leu Ser Ser Lys Ser Val Asn Ala Gln Val Thr Asp Ile Asn Ser
 20 25 30

Lys Gly Leu Glu Leu Arg Lys Thr Val Thr Thr Val Glu Thr Gln Asn
 35 40 45

Leu Glu Gly Leu His His Asp Gly Gln Phe Cys His Lys Pro Cys Pro
 50 55 60

Pro Gly Glu Arg Lys Ala Arg Asp Cys Thr Val Asn Gly Asp Glu Pro
 65 70 75 80

Asp Cys Val Pro Cys Gln Glu Gly Lys Glu Tyr Thr Asp Lys Ala His
 85 90 95

Phe Ser Ser Lys Cys Arg Arg Cys Arg Leu Cys Asp Glu Gly His Gly
 100 105 110

Leu Glu Val Glu Ile Asn Cys Thr Arg Thr Gln Asn Thr Lys Cys Arg
 115 120 125

Cys Lys Pro Asn Phe Phe Cys Asn Ser Thr Val Cys Glu His Cys Asp
 130 135 140

Pro Cys Thr Lys Cys Glu His Gly Ile Ile Lys Glu Cys Thr Leu Thr
 145 150 155 160

Ser Asn Thr Lys Cys Lys Glu Glu Gly Ser Arg Ser Asn Leu Gly Trp
 165 170 175

Leu Cys Leu Leu Leu Pro Ile Pro Leu Ile Val Trp Val Lys Arg
 180 185 190

Lys Glu Val Gln Lys Thr Cys Arg Lys His Arg Lys Glu Asn Gln Gly
 195 200 205

Ser His Glu Ser Pro Thr Leu Asn Pro Glu Thr Val Ala Ile Asn Leu
 210 215 220

Ser Asp Val Asp Leu Ser Lys Tyr Ile Thr Thr Ile Ala Gly Val Met
 225 230 235 240

Thr Leu Ser Gln Val Lys Gly Phe Val Arg Lys Asn Gly Val Asn Glu
 245 250 255

Ala Lys Ile Asp Glu Ile Lys Asn Asp Asn Val Gln Asp Thr Ala Glu
 260 265 270

Gln Lys Val Gln Leu Leu Arg Asn Trp His Gln Leu His Gly Lys Lys
 275 280 285

Glu Ala Tyr Asp Thr Leu Ile Lys Asp Leu Lys Lys Ala Asn Leu Cys
 290 295 300

Thr Leu Ala Glu Lys Ile Gln Thr Ile Ile Leu Lys Asp Ile Thr Ser
 305 310 315 320

Asp Ser Glu Asn Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val Met

325

330

335

Leu Gly Ile Trp Thr Leu Leu Pro Leu Val Leu Thr Ser Val Ala Arg
 340 345 350

Leu Ser Ser Lys Ser Val Asn Ala Gln Val Thr Asp Ile Asn Ser Lys
 355 360 365

Gly Leu Glu Leu Arg Lys Thr Val Thr Thr Val Glu Thr Gln Asn Leu
 370 375 380

Glu Gly Leu His His Asp Gly Gln Phe Cys His Lys Pro Cys Pro Pro
 385 390 395 400

Gly Glu Arg Lys Ala Arg Asp Cys Thr Val Asn Gly Asp Glu Pro Asp
 405 410 415

Cys Val Pro Cys Gln Glu Gly Lys Glu Tyr Thr Asp Lys Ala His Phe
 420 425 430

Ser Ser Lys Cys Arg Arg Cys Arg Leu Cys Asp Glu Gly His Gly Leu
 435 440 445

Glu Val Glu Ile Asn Cys Thr Arg Thr Gln Asn Thr Lys Cys Arg Cys
 450 455 460

Lys Pro Asn Phe Phe Cys Asn Ser Thr Val Cys Glu His Cys Asp Pro
 465 470 475 480

Cys Thr Lys Cys Glu His Gly Ile Ile Lys Glu Cys Thr Leu Thr Ser
 485 490 495

Asn Thr Lys Cys Lys Glu Glu Gly Ser Arg Ser Asn Leu Gly Trp Leu
 500 505 510

Cys Leu Leu Leu Leu Pro Ile Pro Leu Ile Val Val Lys Arg Lys Glu
 515 520 525

Val Gln Lys Thr Cys Arg Lys His Arg Lys Glu Asn Gln Gly Ser His
 530 535 540

Glu Ser Pro Thr Leu Asn Pro Glu Thr Val Ala Ile Asn Leu Ser Asp
 545 550 555 560

Val Asp Leu Ser Lys Tyr Ile Thr Thr Ile Ala Gly Val Met Thr Leu
 565 570 575

Ser Gln Val Lys Gly Phe Val Arg Lys Asn Gly Val Asn Glu Ala Lys
 580 585 590

Ile Asp Glu Ile Lys Asn Asp Asn Val Gln Asp Thr Ala Glu Gln Lys
 595 600 605

Val Gln Leu Leu Arg Asn Trp His Gln Leu His Gly Lys Lys Glu Ala
 610 615 620

Tyr Asp Thr Leu Ile Lys Asp Leu Lys Lys Ala Asn Leu Cys Thr Leu
 625 630 635 640

Ala Glu Lys Ile Gln Thr Ile Ile Leu Lys Asp Ile Thr Ser Asp Ser
 645 650 655

Glu Asn Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val
 660 665

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 909 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Leu Ser Thr Val Pro Asp Leu Leu Pro Leu Val Leu Leu Glu
 1 5 10 15

Leu Leu Val Gly Ile Tyr Pro Ser Gly Val Ile Gly Leu Val Pro His
 20 25 30

Leu Gly Asp Arg Glu Lys Arg Asp Ser Val Cys Pro Gln Gly Lys Tyr
 35 40 45

Ile His Pro Gln Asn Asn Ser Ile Cys Cys Thr Lys Cys His Lys Gly
 50 55 60

Thr Tyr Leu Tyr Asn Asp Cys Pro Gly Pro Gly Asp Thr Asp Cys Arg
 65 70 75 80

Glu Cys Glu Ser Gly Ser Phe Thr Ala Ser Glu Asn His Leu Arg His
 85 90 95

Cys Leu Ser Cys Ser Lys Cys Arg Lys Glu Met Gly Gln Val Glu Ile
 100 105 110

Ser Ser Cys Thr Val Asp Arg Asp Thr Val Cys Gly Cys Arg Lys Asn
 115 120 125

Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe Asn Cys
 130 135 140

Ser Leu Cys Leu Asn Gly Thr Val His Leu Ser Cys Gln Glu Lys Gln
 145 150 155 160

Asn Thr Val Cys Thr Cys His Ala Gly Phe Phe Leu Arg Glu Asn Glu
 165 170 175

 Cys Val Ser Cys Ser Asn Cys Lys Lys Ser Leu Glu Cys Thr Lys Leu
 180 185 190

 Cys Leu Pro Gln Ile Glu Asn Val Lys Gly Thr Glu Asp Ser Gly Thr
 195 200 205

 Thr Val Leu Leu Pro Leu Val Ile Phe Phe Gly Leu Cys Leu Leu Ser
 210 215 220

 Leu Leu Phe Ile Gly Leu Met Tyr Arg Tyr Gln Arg Trp Lys Ser Lys
 225 230 240

 Leu Tyr Ser Ile Val Cys Gly Lys Ser Thr Pro Glu Lys Glu Gly Glu
 245 250 255

 Leu Glu Gly Thr Thr Lys Pro Leu Ala Pro Asn Pro Ser Phe Ser
 260 265 270

 Pro Thr Pro Gly Phe Thr Pro Thr Leu Gly Phe Ser Pro Val Pro Ser
 275 280 285

 Ser Thr Phe Thr Ser Ser Ser Thr Tyr Thr Pro Gly Asp Cys Pro Asn
 290 295 300

 Phe Ala Ala Pro Arg Arg Glu Val Ala Pro Pro Tyr Gln Gly Ala Asp
 305 310 320

 Pro Ile Leu Ala Thr Ala Leu Ala Ser Asp Pro Ile Pro Asn Pro Leu
 325 330 335

 Gln Lys Trp Glu Asp Ser Ala His Lys Pro Gln Ser Leu Asp Thr Asp
 340 345 350

 Asp Pro Ala Thr Leu Tyr Ala Val Val Glu Asn Val Pro Pro Leu Arg
 355 360 365

 Trp Lys Glu Phe Val Arg Arg Leu Gly Leu Ser Asp His Glu Ile Asp
 370 375 380

 Arg Leu Glu Leu Gln Asn Gly Arg Cys Leu Arg Glu Ala Gln Tyr Ser
 385 390 395 400

 Met Leu Ala Thr Trp Arg Arg Arg Thr Pro Arg Arg Glu Ala Thr Leu
 405 410 415

 Glu Leu Leu Gly Arg Val Leu Arg Asp Met Asp Leu Leu Gly Cys Leu
 420 425 430

 Glu Asp Ile Glu Glu Ala Leu Cys Gly Pro Ala Ala Leu Pro Pro Ala
 435 440 445

 Pro Ser Leu Leu Arg Met Gly Leu Ser Thr Val Pro Asp Leu Leu Leu

450	455	460
Pro Leu Val Leu Leu Glu Leu Leu Val Gly Ile Tyr Pro Ser Gly Val		
465	470	475
Ile Gly Leu Val Pro His Leu Gly Asp Arg Glu Lys Arg Asp Ser Val		
485		490
Cys Pro Gln Gly Lys Tyr Ile His Pro Gln Asn Asn Ser Ile Cys Cys		
500	505	510
Thr Lys Cys His Lys Gly Thr Tyr Leu Tyr Asn Asp Cys Pro Gly Pro		
515	520	525
Gly Gln Asp Thr Asp Cys Arg Glu Cys Ser Gly Ser Phe Thr Ala		
530	535	540
Ser Glu Asn His Leu Arg His Cys Leu Ser Cys Ser Lys Cys Arg Glu		
545	550	555
Lys Glu Met Gly Gln Val Glu Ile Ser Ser Cys Thr Val Asp Arg Asp		
565	570	575
Thr Val Cys Gly Cys Arg Lys Asn Gln Tyr Arg His Tyr Trp Ser Glu		
580	585	590
Asn Leu Phe Gln Cys Phe Asn Cys Ser Leu Cys Leu Asn Gly Thr Val		
595	600	605
His Leu Ser Cys Gln Glu Lys Gln Asn Thr Val Cys Thr Cys His Ala		
610	615	620
Gly Phe Phe Leu Arg Glu Asn Glu Cys Val Ser Cys Ser Asn Cys Lys		
625	630	635
Lys Ser Leu Glu Cys Thr Lys Leu Cys Leu Pro Gln Ile Glu Asn Val		
645	650	655
Lys Gly Thr Glu Asp Ser Gly Thr Thr Val Leu Leu Pro Leu Val Ile		
660	665	670
Phe Phe Gly Leu Cys Leu Leu Ser Leu Leu Phe Ile Gly Leu Met Tyr		
675	680	685
Arg Tyr Gln Arg Trp Lys Ser Asp Leu Tyr Ser Ile Val Cys Gly Lys		
690	695	700
Ser Thr Pro Glu Lys Glu Gly Glu Leu Glu Gly Thr Thr Lys Pro		
705	710	715
Leu Ala Pro Asn Pro Ser Phe Ser Pro Thr Pro Gly Phe Thr Pro Thr		
725	730	735
Leu Gly Phe Ser Pro Val Pro Ser Ser Thr Phe Thr Ser Ser Ser Thr		
740	745	750

Tyr Thr Pro Gly Asp Cys Pro Asn Phe Ala Ala Pro Arg Arg Glu Val
 755 760 765
 Ala Pro Pro Tyr Gln Gly Ala Asp Pro Ile Leu Ala Thr Ala Leu Ala
 770 775 780
 Ser Asp Pro Ile Pro Asn Pro Leu Gln Lys Trp Glu Asp Ser Ala His
 785 790 795 800
 Lys Pro Gln Ser Leu Asp Thr Asp Asp Pro Ala Thr Leu Tyr Ala Val
 805 810 815
 Val Glu Asn Val Pro Pro Leu Arg Trp Lys Glu Phe Val Arg Arg Leu
 820 825 830
 Gly Leu Ser Pro His Glu Ile Asp Arg Leu Glu Leu Gln Asn Gly Arg
 835 840 845
 Cys Leu Arg Glu Ala Gln Tyr Ser Met Leu Ala Thr Trp Arg Arg Arg
 850 855 860
 Thr Pro Arg Arg Glu Ala Thr Leu Glu Leu Leu Gly Arg Val Leu Arg
 865 870 875 880
 Asp Met Asp Leu Leu Gly Cys Leu Glu Asp Ile Glu Glu Ala Leu Cys
 885 890 895
 Gly Pro Ala Ala Leu Pro Pro Ala Pro Ser Leu Leu Arg
 900 905

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 833 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Glu Gln Arg Pro Arg Gly Cys Ala Ala Val Ala Ala Leu Leu
 1 5 10 15
 Leu Val Leu Leu Gly Ala Arg Ala Gln Gly Gly Thr Arg Ser Pro Arg
 20 25 30
 Cys Asp Cys Ala Gly Asp Phe His Lys Lys Ile Gly Leu Phe Cys Cys
 35 40 45

Arg Gly Cys Pro Ala Gly His Tyr Leu Lys Ala Pro Cys Thr Glu Pro
 50 55 60
 Cys Gly Asn Ser Thr Cys Leu Val Cys Pro Gln Asp Thr Phe Leu Ala
 65 70 75 80
 Trp Glu Asn His His Asn Ser Glu Cys Ala Arg Cys Gln Ala Cys Asp
 85 90 95
 Glu Gln Ala Ser Gln Val Ala Leu Glu Asn Cys Ser Ala Val Ala Asp
 100 105 110
 Thr Arg Cys Gly Cys Lys Pro Gly Trp Phe Val Glu Cys Gln Val Ser
 115 120 125
 Gln Cys Val Ser Ser Ser Pro Phe Tyr Cys Gln Pro Cys Leu Asp Cys
 130 135 140
 Gly Ala Leu His Arg His Thr Arg Leu Leu Cys Ser Arg Arg Asp Thr
 145 150 155 160
 Asp Cys Gly Thr Cys Leu Pro Gly Phe Tyr Glu His Gly Asp Gly Cys
 165 170 175
 Val Ser Cys Pro Thr Ser Thr Leu Gly Ser Cys Pro Glu Arg Cys Ala
 180 185 190
 Ala Val Cys Gly Trp Arg Gln Met Phe Trp Val Gln Val Leu Leu Ala
 195 200 205
 Gly Leu Val Val Pro Leu Leu Leu Gly Ala Thr Leu Thr Tyr Thr Tyr
 210 215 220
 Arg His Cys Trp Pro His Lys Pro Leu Val Thr Ala Asp Glu Ala Gly
 225 230 235 240
 Met Glu Ala Leu Thr Pro Pro Pro Ala Thr His Leu Ser Pro Leu Asp
 245 250 255
 Ser Ala His Thr Leu Leu Ala Pro Pro Asp Ser Ser Glu Lys Ile Cys
 260 265 270
 Thr Val Gln Leu Val Gly Asn Ser Trp Thr Pro Gly Tyr Pro Glu Thr
 275 , 280 285
 Gln Glu Ala Leu Cys Pro Gln Val Thr Trp Ser Trp Asp Gln Leu Pro
 290 295 300
 Ser Arg Ala Leu Gly Pro Ala Ala Ala Pro Thr Leu Ser Pro Glu Ser
 305 310 315 320
 Pro Ala Gly Ser Pro Ala Met Met Leu Gln Pro Gly Pro Gln Leu Tyr
 325 330 335
 Asp Val Met Asp Ala Val Pro Ala Arg Arg Trp Lys Glu Phe Val Arg

340

345

350

Thr Leu Gly Leu Arg Glu Ala Glu Ile Glu Ala Val Glu Val Glu Ile
355 360 365

Gly Arg Phe Arg Asp Gln Gln Tyr Glu Met Leu Lys Arg Trp Arg Gln
370 375 380

Gln Gln Pro Ala Gly Leu Gly Ala Val Tyr Ala Ala Leu Glu Arg Met
385 390 395 400

Gly Leu Asp Gly Cys Val Glu Asp Leu Arg Ser Arg Leu Gln Arg Gly
405 410 415

Pro Met Glu Gln Arg Pro Arg Gly Cys Ala Ala Val Ala Ala Ala Leu
420 425 430

Leu Leu Val Leu Leu Gly Ala Arg Ala Gln Gly Gly Thr Arg Ser Pro
435 440 445

Arg Cys Asp Cys Ala Gly Asp Phe His Lys Lys Ile Gly Leu Phe Cys
450 455 460

Cys Arg Gly Cys Pro Ala Gly His Tyr Leu Lys Ala Pro Cys Thr Glu
 465 470 475 480

Ala Trp Glu Asn His His Asn Ser Glu Cys Ala Arg Cys Gln Ala Cys
500 505 510

Asp Glu Ala Ser Gln Val Ala Leu Glu Asn Cys Ser Ala Val Ala Asp
515 520 525

Thr Arg Cys Gly Cys Lys Pro Gly Trp Phe Val Glu Cys Gln Val Ser
 530 535 540

Gly Ala Leu His Arg His Thr Arg Leu Leu Cys Ser Arg Arg Asp Thr
565 570 575

Asp Cys Gly Thr Cys Leu Pro Gly Phe Tyr Glu His Gly Asp Gly Cys
580 585 590

Val Ser Cys Pro Thr Ser Thr Leu Gly Ser Cys Pro Glu Arg Cys Ala
595 600 605

Ala Val Cys Gly Trp Arg Gln Met Phe Trp Val Gln Val Leu Leu Ala
610 615 620

Gly Leu Val Val Pro Leu Leu Leu Gly Ala Thr Leu Thr Tyr Thr Tyr
 625 630 635 640

Arg His Cys Trp Pro His Lys Pro Leu Val Thr Ala Asp Glu Ala Gly
 645 650 655
 Met Glu Ala Leu Thr Pro Pro Ala Thr His Leu Ser Pro Leu Asp
 660 665 670
 Ser Ala His Thr Leu Leu Ala Pro Pro Asp Ser Ser Glu Lys Ile Cys
 675 680 685
 Thr Val Gln Leu Val Gly Asn Ser Trp Thr Pro Gly Tyr Pro Glu Thr
 690 695 700
 Gln Glu Ala Leu Cys Pro Gln Val Thr Trp Ser Trp Asp Gln Leu Pro
 705 710 715 720
 Ser Arg Ala Leu Gly Pro Ala Ala Ala Pro Thr Leu Ser Pro Glu Ser
 725 730 735
 Pro Ala Gly Ser Pro Ala Met Met Leu Gln Pro Gly Pro Gln Leu Tyr
 740 745 750
 Asp Val Met Asp Ala Val Pro Ala Arg Arg Trp Lys Glu Phe Val Arg
 755 760 765
 Thr Leu Gly Leu Arg Glu Ala Glu Ile Glu Ala Val Glu Val Glu Ile
 770 775 780
 Gly Arg Phe Arg Asp Gln Gln Tyr Glu Met Leu Lys Arg Trp Arg Gln
 785 790 795 800
 Gln Gln Pro Ala Gly Leu Gly Ala Val Tyr Ala Ala Leu Glu Arg Met
 805 810 815
 Gly Leu Asp Gly Cys Val Glu Asp Leu Arg Ser Arg Leu Gln Arg Gly
 820 825 830
 Pro

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 426 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGCANAGGTN CGTACCTAGC TCACCTGCAA CCATCAAAC T ⁶⁰
CACAGCAATG GGAAACATAG CCCTTGAA GANTGTNTC CACCAGGATC T ¹²⁰
AAACATCCTG GGAGCCTGTT AACCGGTGCC CCAAAGGNTG GTCAAGGTCA AGGAATIGTT ¹⁸⁰
NCGCCCTGGA AGTGAACATC GAGTGTNTCC ACAAAGGATT CAGGCAATGG GACATAAATA ²⁴⁰
TATGGGTGAA TTTTGGTTGT GAACTTTGGT TGNTCCCGTT GNTGTTGNTG GCTGTGCTGA ³⁰⁰
TTGTTTGTGTTG TTGCATCGGC TTCAGGTTNT GGAGGGGAC CCAAGTGCAT GGACAGGGTG ³⁶⁰
TGTTTCTGGG GTTTGGGTCT CTTAGAGGGC NTGGGTTANG GCANGTTCAC AAGGGTTTA ⁴²⁰
GCAANG ⁴²⁶

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 339 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGGGGCTGAG GACAATGCTG ACNACGAGAT TCTGAGCAAC GCAGNACTNG CTGTCCACTT ⁶⁰
TCGTCTNTGN GCAGCAAATG GAAAGCCAGG AGCCGGCAGA TTTGACAGGT GTCACTGTAC ¹²⁰
AGTCCCCAGG GGAGGCACAG TGTCTGCTGG TGAGTTGGGG ACAGGCCCTT GCAAGACCTT ¹⁸⁰
GTGAGGCAGG GGGTGAAGGC CATGNCTCGG CTTCNNTGG TCAAAGGGGA AGTGGAGCCT ²⁴⁰
GAGGGAGATG GGACTTNAGG GGGACGGNGC TGCCTGGGG AAAAGCAGCC ACCNTTTGAC ³⁰⁰
AAGGGGGACA GGCATTTTN CAAATGTGTG CTTNTTGGT ³³⁹

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGGCATGCA TGATCAATCA ATTGGCAC

28

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCGGGATCCG CCATCATGGC GCCACCACCA GCTAGA

36

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGGGATCCT CACTCCAAGG ACACGGCAGA GCC

33

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCGGGATCCT CAATTATGTC CATTGCCTG

29